Check for updates

JGR | Machine Learning and Computation

RESEARCH ARTICLE
10.1029/2024JH000292

Automatic Detection and Classification of Aurora in THEMIS All-Sky Images

Jeremiah W. Johnson[1], Doğacan Su Öztürk[2], Donald Hampton[2], Hyunju K. Connor[3], Matthew Blandin[2], and Amy Keesee[4,5]

[1]Department of Applied Engineering & Sciences, University of New Hampshire, Manchester, NH, USA, [2]The Geophysical Institute, University of Alaska-Fairbanks, Fairbanks, AK, USA, [3]NASA Goddard Space Flight Center, Greenbelt, MD, USA, [4]Department of Physics & Astronomy, University of New Hampshire, Durham, NH, USA, [5]Institute for the Study of Earth, Oceans and Space, University of New Hampshire, Durham, NH, USA

Key Points:

- We describe a novel machine-learning algorithm for classifying THEMIS white-light all–sky images into one of six distinct categories
- We analyze classification results in conjunction with ground magnetometer data for the years 2008–2022
- The publicly available data set will enable extensive statistical studies and machine-learning analyses of auroral labels

Correspondence to:
J. W. Johnson,
jeremiah.johnson@unh.edu

Citation:
Johnson, J. W., Öztürk, D. S., Hampton, D., Connor, H. K., Blandin, M., & Keesee, A. (2024). Automatic detection and classification of aurora in THEMIS all-sky images. *Journal of Geophysical Research: Machine Learning and Computation*, *1*, e2024JH000292. https://doi.org/10.1029/2024JH000292

Received 30 MAY 2024
Accepted 10 NOV 2024

Author Contributions:
Conceptualization: Jeremiah W. Johnson, Doğacan Su Öztürk, Donald Hampton, Hyunju K. Connor, Matthew Blandin
Data curation: Jeremiah W. Johnson, Doğacan Su Öztürk
Formal analysis: Jeremiah W. Johnson, Doğacan Su Öztürk
Funding acquisition: Jeremiah W. Johnson, Doğacan Su Öztürk, Donald Hampton, Hyunju K. Connor
Investigation: Jeremiah W. Johnson, Doğacan Su Öztürk
Methodology: Jeremiah W. Johnson, Doğacan Su Öztürk, Donald Hampton, Hyunju K. Connor
Project administration: Jeremiah W. Johnson

© 2024 The Author(s). Journal of Geophysical Research: Machine Learning and Computation published by Wiley Periodicals LLC on behalf of American Geophysical Union.
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Abstract We report a novel machine-learning algorithm for automatically detecting and classifying aurora in all–sky images (ASI) that is largely trained without requiring ground–truth labels. By including a small number of labeled images, we are able to automatically label all of the approximately 700 million images in the Time History of Events and Macroscale Interactions during Substorms (THEMIS) ASI data set from 2008 to 2022. We use a two–stage approach. In the first stage, we adapt the Simple framework for Contrastive Learning of Representations (SimCLR) algorithm to learn latent representations of THEMIS all–sky images. We then finetune a classifier network on the latent representations our model learns of the manually labeled Oslo aurora THEMIS (OATH) data set. We demonstrate that this two–stage approach achieves excellent classification results on data for which there is no current ML classification benchmark. The outcome of this work will facilitate efficient information retrieval for researchers interested in specific categories of aurora and will enable large scale statistical studies and machine learning analyses of THEMIS all–sky images that have not previously been possible. To demonstrate possible ways to utilize this database, we performed a statistical analysis of the occurrence rates of auroral labels with respect to solar wind parameters, interplanetary magnetic field vector, and geomagnetic indices. We further investigate the occurrence rates of auroral phenomena in the annotated data set and their geoeffectiveness by utilizing the co–located THEMIS ground magnetometer data set.

Plain Language Summary A vast number of white–light all sky images have been collected to date by ground–based all–sky imagers distributed across North America through the Time History of Events and Macroscale Interactions during Substorms (THEMIS) mission. This massive data set is a valuable resource that can aid researchers in understanding the processes that lead to different auroral forms, but its size limits how effectively and efficiently the data can be filtered, sorted and retrieved. A first step toward efficient information retrieval that will enable large–scale statistical studies and machine–learning driven analyses is to coarsely annotate the images to identify and describe any auroral phenomena present. In this work, we describe a novel machine learning algorithm that efficiently annotates THEMIS all-sky images (ASI) using six distinct categories: arc, diffuse, discrete, cloudy, moon, and clear/no aurora, and we summarize classification results and associated ground magnetic field signatures for years 2008–2022.

1. Introduction

Obtaining a better understanding of the interaction of the solar wind with the Earth's magnetosphere and upper atmosphere is a primary goal of recent heliophysics research (NASA Science Mission Directorate (SMD) 2014 Science Plan, n.d.; National Research Council, 2013). Aurora borealis and aurora australis are formed when the coupling of the solar wind and the magnetosphere causes high–energy electrons and ions to precipitate from the magnetosphere to the upper atmosphere. Auroral features at a scale of 10–10,000 km hold key information concerning magnetospheric coupling processes at a scale of 1–100 Earth radii. Given that there are only a limited number of satellites available to monitor the entire magnetosphere, ground–based auroral images have been critical in improving understanding of the propagation of solar wind energy to near–Earth space (Akasofu, 1964; Akasofu, 1981; L. Clausen et al., 2014).

With decades of ground and space observations, global aurora patterns at the 1,000–10,000 km scale and their general mechanisms are well understood (Newell et al., 2009). However, the morphology of smaller–scale aurora

29935210, 2024, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024JH000292 by Nasa Goddard Space Flight, Wiley Online Library on [10/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Software: Jeremiah W. Johnson, Doğacan Su Öztürk
Validation: Jeremiah W. Johnson, Doğacan Su Öztürk
Visualization: Jeremiah W. Johnson, Doğacan Su Öztürk
Writing – original draft: Jeremiah W. Johnson, Doğacan Su Öztürk
Writing – review & editing: Jeremiah W. Johnson, Doğacan Su Öztürk

forms (less than 1000 km) and their connections to mid– to large–scale magnetospheric dynamics (less than 100 Earth radii) are far less well understood. A better understanding of smaller–scale auroral morphology is critical to improving our understanding of the solar wind–magnetosphere–upper atmosphere interaction.

The heliophysics community has amassed massive collections of images of aurora borealis and aurora australis; indeed, the sheer abundance of auroral imagery itself presents a challenge. The vast majority of this data is publicly available but unlabeled or not organized in a way that is amenable to large–scale statistical studies or supervised machine learning. Coarsely classifying this data is a first and important key step toward identifying local–scale auroral features and enabling deeper analyses of the link between auroral morphology and magnetospheric dynamics.

1.1. Related Work

Automatic classification of auroral images has been studied for several decades. Early approaches utilizing traditional computer vision techniques reliant on the construction of hand–designed features include Syrjäsuo and Donovan (2002); Syrjäsuo and Donovan (2004); Syrjäsuo et al. (2007); Q. Yang et al. (2012); Rao et al. (2014), with the best of these able to identify four distinct categories of aurora with a positive detection rate of up to 85%.

The field of computer vision has for the past decade or more been dominated by deep learning algorithms. Convolutional neural networks and their variants in particular have achieved remarkable performance on a wide range of challenging tasks (He et al., 2016; Krizhevsky et al., 2012; LeCun et al., 1998). There have been several recent efforts to use deep neural networks for automatic auroral classification. In L. B. Clausen and Nickisch (2018a), 5,824 images of aurora obtained from THEMIS all–sky imagers are manually assigned one of the distinct categories *Arc*, *Discrete*, *Diffuse*, *Clear/No Aurora*, *Cloudy*, and *Moon*. Using an Inception model (Szegedy et al., 2015) pretrained on the ImageNet data set (Deng et al., 2009; Sharif Razavian et al., 2014a), they extract feature vectors and fit a ridge classifier to them to obtain an 82% accuracy. This approach was built upon in Sado et al. (2022, 2023) to predict the onset of auroral substorms. In Kvammen et al. (2020), the authors manually classified 3,846 auroral images into the categories *Arc aurora*, *Auroral breakup*, *Colored aurora*, *Discrete aurora*, *Edge aurora*, *Faint Aurora*, *Patchy Aurora*, then compare classification results from a range of machine learning algorithms, including convolutional neural network (CNN)–based architectures. In Nanjo et al. (2022), a deep CNN–based ResNet50 (He et al., 2016) architecture is used to identify auroral activity in ASI from Tromsø, Norway. And in Johnson et al. (2021), a SimCLR model is used to learn latent representations of the Oslo Aurora THEMIS data set that are then classified using a simple linear classification algorithm, thus demonstrating that a SimCLR–based model may be a suitable approach for auroral image analysis on a larger scale.

There are two closely related challenges inherent to using deep neural networks for auroral image analysis. First, neural networks are usually trained using a supervised learning method. Supervised learning requires a substantial quantity of unambiguous, high–quality ground–truth labeled data for training. While there exist massive amounts of auroral image data, the quality of this data varies dramatically and to date, very little of it has been labeled. The use of ImageNet–pretrained convolutional neural networks (CNNs) as in (L. B. Clausen & Nickisch, 2018a) presents a possible workaround to this issue as pretraining is often effective at reducing the amount of data necessary to train a CNN to convergence (Sharif Razavian et al., 2014b), but pretraining only partially addresses the challenge of limited labeled data and may introduce other issues and biases (Geirhos et al., 2018). Semi–supervised learning presents another possible workaround. In a semi–supervised framework, a model is first trained on a small manually annotated subset of a larger unlabeled data set. Once the model's performance is satisfactory on the labeled data, its high–confidence predictions on labeled data are treated as ground truth annotations, and these data are added to the training data set. This process is then iterated until the full data set is annotated. As with ImageNet pretraining, this approach holds promise but can introduce other issues and biases, particularly when the ratio of labeled to unlabeled images in the data set is small (Kim et al., 2020; X. Yang et al., 2022).

Second and more importantly, manually labeling auroral images itself presents additional challenges beyond the time and expense required: labels for auroral images are often chosen based on subjective, qualitative judgments, and there is a frequent disagreement between researchers regarding what an appropriate label set is. Complicating the picture further is the fact that even after a label set is decided on it is common for all–sky auroral images to capture more than one type of auroral phenomenon in the same frame, raising the additional question of how to decide what the correct classification should be in these ambiguous cases.

29935210, 2024, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024JH000292 by Nasa Goddard Space Flight, Wiley Online Library on [10/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Self–supervised learning is a recently developed training strategy for deep neural networks that does not require any ground–truth labels. Instead, a synthetic proxy training objective is formulated such that the labels for the proxy training objective can be automatically constructed; then, the network is trained with these labels using a conventional supervised algorithm. Post–training, the proxy labels are discarded and the representations that the model learns can be extracted and used for downstream tasks. In the past few years, self–supervised pretraining followed by classification with simple linear classifiers (Caron et al., 2020; Chen et al., 2020; Dai et al., 2021) has been shown to achieve levels of performance rivaling those of conventional supervised algorithms on benchmark machine learning tasks and data sets, and a similar approach in auroral image analysis has shown promise (Johnson et al., 2021; Partamies et al., 2024). Self–supervision directly address the challenges just described in auroral imaging simultaneously: the learned representations of auroral images may be useful for quantitatively determining meaningful labels, and the technique can be used to apply such labels automatically, facilitating the creation of much larger labeled auroral image data sets. Our contribution in this work is as follows:

- We adapt the *Simple Framework for Contrastive Learning of Representations* (SIMCLR) model introduced in Chen et al. (2020) to the THEMIS all–sky image (ASI) database to obtain representations of approximately 700 million auroral images collected from 2008 to 2022 (Donovan et al., 2006). Using a small manually annotated sample of this database released with the publication of L. B. Clausen and Nickisch (2018a), we then train a classifier to manually categorize the ASI into the distinct categories *Arc*, *Discrete*, *Diffuse*, *Clear/No Aurora*, *Cloudy*, and *Moon*.
- We provide a detailed analysis of the drivers and impacts of the different auroral types predicted by the model for the years 2008–2022.

The rest of the paper is organized as follows: in Section 2, we describe our model and training strategy. In Section 3, we discuss the annotations obtained with our model. In Section 4, we analyze the drivers and impacts of different auroral types predicted by the model for the years 2008–2022. In Section 5, we conclude with a discussion of future work.

2. Methodology

2.1. Data

We utilize two overlapping image data sets for the training of the model. The first is the data set consisting of all–sky images collected as part of the Time History of Events and Macroscale Interactions during Substorms (THEMIS) mission. This data set is comprised of white–light all–sky images collected at stations across North America 2006–present. For this study, we rely on the approximately 700 million images in this data set collected from 2008 to 2022. The locations of the THEMIS all–sky imagers can be found in Figure 3 of the Donovan et al. (2006). The all–sky imagers provide broadband images with a frame rate of $256x256x16 - bit$ every 3 s. The second image data set used in this study consists of 5,824 images manually curated from various THEMIS all–sky imagers and assigned one of the labels arc, diffuse, discrete, cloudy, moon, and clear. These labels are precisely defined below. This data set was initially constructed for the study presented in L. B. Clausen and Nickisch (2018a), and is publicly available at http://tid.uio.no/plasma/oath/oath_v1.1_20181026.tgz. We refer to this data set as the Oslo Auroral THEMIS (OATH) data set.

Each image $\boldsymbol{x}$ used in L. B. Clausen and Nickisch (2018a) was circularly cropped by 15% to remove pixels corresponding to very low elevation angles and then scaled to enhance dim features by applying the following formula:

$$x_{ij} = \max_{i,j}\left(\min_{i,j}\left(\frac{x_{ij} - m_{\boldsymbol{x}}}{M_{\tilde{\boldsymbol{x}}}}, 1\right), 0\right), \tag{1}$$

where x_{ij} denotes the $i,j - th$ pixel value, $m_{\boldsymbol{x}}$ denotes the 1^{st} percentile brightness value, and $M_{\tilde{\boldsymbol{x}}}$ denotes the 99^{th} percentile brightness of $\boldsymbol{x} - m_{\boldsymbol{x}}$. The resulting cropped and scaled images in the OATH data set were manually assigned a distinct label from the categories *Arc* (774 images), *Discrete* (1,102 images), *Diffuse* (1,400 images), *Cloudy* (817 images), *Moon* (585 images), *Clear/No aurora* (1,082 images). These categories are described in L. B. Clausen and Nickisch (2018a) as follows and examples from each class are displayed in Figure 1:

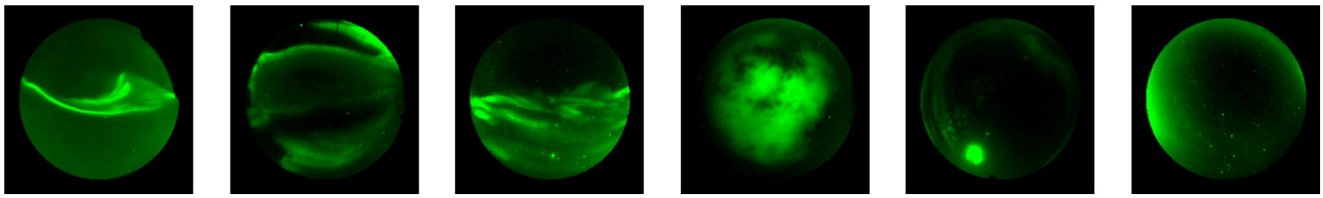


Figure 1. A sample of images from the Oslo Aurora THEMIS data set (OATH). From left to right, the manually assigned ground–truth labels are "Arc," "Diffuse," "Discrete," "Cloudy," "Moon," "Clear." Note that these images are false color images; they have been converted to RGB with intensity information recorded in the green channel. This is the required data type for the model used in (L. B. Clausen & Nickisch, 2018a), and we follow that format for consistency.

- *Arc*: Showing one or multiple bands of aurora that stretch across the field-of-view; typically, the arcs have well-defined, sharp edges.
- *Diffuse*: Large patches of aurora, typically with fuzzy edges.
- *Discrete*: Auroral forms with well-defined, sharp edges, that are, however, not arc–like.
- *Cloudy*: Dominated by clouds or the dome of the imager is covered with snow.
- *Moon*: Dominated by light from the Moon.
- *Clear/No aurora*: Images which show a clear sky (stars and planets are clearly visible) without the appearance of aurora.

As noted in (L. B. Clausen & Nickisch, 2018a), there is a substantial overlap between the chosen categories and those used in other studies, but they are not often an exact match; moreover, these categories are not mutually exclusive and the correct label to assign may at times be ambiguous.

For the statistical analysis, we use the 1-min averages of the Interplanetary Magnetic Field (IMF) vector, IMF magnitude, clock angle, solar wind velocity, density, temperature, dynamic pressure, solar wind electric field, solar wind Mach number, symH index, and auroral electrojet (AE) data from the OMNI database (https://omniweb.gsfc.nasa.gov/). We also utilize the level-22 magnetic perturbation data from THEMIS ground magnetometers which are co–located with all–sky imagers (https://themis.ssl.berkeley.edu/data/themis/thg/l2/mag/).

2.2. SimCLR

The Simple framework for Contrastive Learning of Representations (SimCLR) algorithm was introduced in Chen et al. (2020). As the name suggests, SimCLR is a simple yet highly effective algorithm for unsupervised representation learning. The SimCLR framework consists of four components, described below and diagrammed in Figure 2:

- A stochastic *data augmentation module*. Given a sample $\boldsymbol{x}$, the data augmentation module randomly applies two transformations $\mathcal{T}_i, \mathcal{T}_j$ to $\boldsymbol{x}$, resulting in two distinct views $\boldsymbol{x}_i, \boldsymbol{x}_j$ of $\boldsymbol{x}$ that we consider a *positive pair* in the context of the contrastive loss function described below.
- A *base encoder*. The base encoder is used to extract an initial representation of a transformed input $\mathcal{T}_k(\boldsymbol{x})$. In principle, this could be any machine learning model; in most applications, including this one, a residual neural network (ResNet) (He et al., 2016) is used.
- A *projection head*. The projection head maps the output of the base encoder to the space where the contrastive loss function is to be applied. Again, in principle this could be any machine learning model; in practice the projection head is usually a shallow (one–layer) dense neural network.
- A *contrastive loss function* designed for the following contrastive prediction task: given a set of pairs of examples, that includes a positive pair, identify the positive pair.

The contrastive loss function used here is the *normalized temperature–scaled cross–entropy loss* (Oord et al., 2018; Sohn, 2016; Wu et al., 2018). Letting $\text{sim}(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{x}^T\boldsymbol{y}/\|\boldsymbol{x}\|\|\boldsymbol{y}\|$ (*cosine similarity*), the loss for a positive pair of examples $\boldsymbol{x}_i$, $\boldsymbol{x}_j$ is given by

$$\ell(i,j) = -\log \frac{\exp(\text{sim}(z_i,z_j)/\tau)}{\sum_{k=1}^{2N} \mathbb{1}_{[k\neq i]} \exp(\text{sim}(z_k,z_i)/\tau)}. \tag{2}$$

29935210, 2024, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024JH000292 by Nasa Goddard Space Flight, Wiley Online Library on [10/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

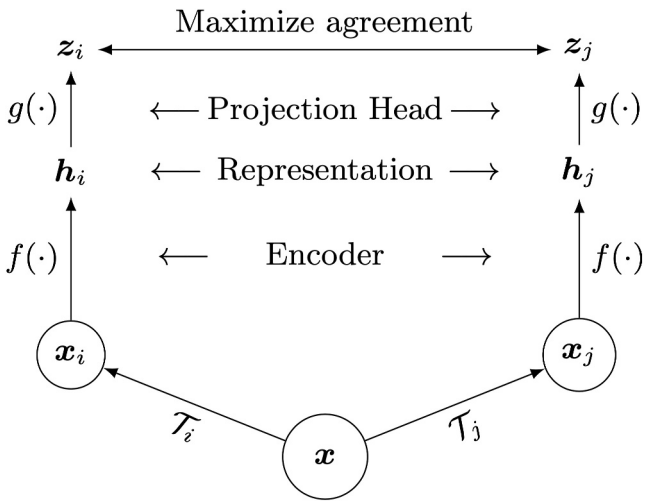


Figure 2. A simple framework for contrastive learning of visual representations. Given a batch of images, for each sample $\boldsymbol{x}$, two distinct transformations are randomly applied ($\mathcal{T}_i$ and $\mathcal{T}_j$) to obtain two correlated views of $\boldsymbol{x}$, a "positive pair." A base encoder network $f(\cdot)$ and a projection head $g(\cdot)$ are then trained to identify the positive pairs among pairs constructed from all transformed samples in a minibatch. After training is completed, the encoder $f(\cdot)$ is applied to each $\boldsymbol{x}$ to obtain a representation $\boldsymbol{h}$ which can then be used for downstream tasks. Diagram adapted from Chen et al. (2020).

In Equation 2, z_i, z_j are the outputs of the projection head for two views $\boldsymbol{x}_i, \boldsymbol{x}_j$ of an input $\boldsymbol{x}$. τ is a temperature parameter, $\mathbb{1}_{[k \neq i]}$ is an indicator function that evaluates to one if $i \neq k$, and the $2N$ examples represented in the sum in the denominator are obtained by using the $2(N-1)$ augmented pairs in a minibatch as negative examples. This mitigates any need to sample negative pairs explicitly. The model is then trained to minimize

$$\frac{1}{2N}\sum_{k=1}^{N}[\ell(2k-1,2k)+\ell(2k,2k-1)]. \tag{3}$$

The choice of transformations to be sampled is believed to be a critical component in learning successful representations using the SimCLR algorithm. The intuition behind the SimCLR model is that just as a human can recognize objects despite a transformation of one's view of them (e.g., recognizing a person despite viewing them in profile with much of their face occluded, or through a stained glass window that changes our perception of the color of their features), so too should the model recognize when it is presented with two differing views of the same phenomenon. The transformations chosen determine the differing views the model receives of the input images. The pre–processing done for the OATH data set and that we follow here for consistency first converts the images to RGB images with intensity data in the green channel. This enables a color distortion transformation where the red, green, and blue components of each pixel are changed while the intensity is maintained. Analogously, random cropping can be used to provide a different view of an input image: if an image contains an arc, for example, then an appropriately sized crop of the image will contain at least a portion of an arc, and the model will learn to recognize that the arc is the aspect of the image that it can use to assign the appropriate label. We incorporate random crops and color distortions as well as random flips over both horizontal and vertical axes motivated by previous research demonstrating that random cropping and color distortions are two transformations that are highly likely to lead to high–quality representation learning (Chen et al., 2020), and the fact that the morphologies captured in the set of annotations used in this study are largely invariant to these transformations.

In all of our experiments, we use a projection head consisting of two fully connected linear layers with an intermediate rectified linear unit (ReLU) activation and without biases; that is, $g(\boldsymbol{h}) = W_2(\max(0, W_1(\boldsymbol{h}))$, where W_1 and W_2 denote weight matrices. As an encoder, we use a lightweight ResNet18 model that is randomly initialized (He et al., 2016). The temperature hyperparameter τ of the loss function (cf. Equation 2) was set to 0.5.

2.3. Self–Supervised Training

To train the SimCLR model, we selected five dates known to have high auroral activity (2014–03–26, 2015–10–07, 2016–10–25, 2021–02–7, and 2021–03–14), and we created a training data set consisting of all of the data available from all of the THEMIS imagers on those dates, 910,155 images in total. The training images were preprocessed by cropping and normalizing them using the same steps described in the preprocessing of the OATH data set (cf. Section 2.1). We then train the SimCLR model on these data for 60 epochs, using the Adam optimization algorithm (Kingma & Ba, 2014) with a fixed learning rate of 0.0003, no weight decay, and a minibatch size of 320. All models were implemented using the open source machine learning library PyTorch (Paszke et al., 2019) and trained using a single NVIDIA Quadro RTX 5000 GPU with 18 GB of onboard memory. Training with these hyperparameter choices on this hardware takes approximately 10 days.

2.4. Finetuning

The SimCLR model produces high quality latent representations that can be used for a variety of downstream tasks, but it is not a classifier on its own without modification. After self–supervised training, we extract the SimCLR encoder ($f(\cdot)$ in Figure 2) and attach a new prediction head to it. This head is identical in architecture to the projection head utilized for self–supervised training ($g(\cdot)$ in Figure 2), but is randomly initialized and is

29935210, 2024, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024JH000292 by Nasa Goddard Space Flight, Wiley Online Library on [10/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

29935210, 2024, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024JH000292 by Nasa Goddard Space Flight, Wiley Online Library on [10/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

followed by a softmax activation that outputs the probability of each label being the correct one to assign to the input image. We then freeze the weights of the encoder for a period of 10 epochs and train the prediction head to classify the OATH data set using full supervision. After the initial 10 epochs of training, we reduce the learning rate, unfreeze the weights of the encoder, and continue supervised training with the OATH data set for 90 more epochs, 100 epochs of training in total. Given the vastly reduced size of the data set used for finetuning, this stage of training was completed in less than an hour. The fully trained model can classify an ASI in a matter of milliseconds.

2.5. Data Analysis

In Section 4, we discuss our findings from the self–supervised model. We also conducted a statistical data analysis using the labels obtained from applying the self–supervised model to THEMIS ASI data in which we study the relationship between the labels and external drivers, such as solar wind and interplanetary magnetic field parameters, as well as geomagnetic activity indices. We further investigate the geoeffectiveness of different aurora types by utilizing the THEMIS ground magnetometers which are co–located with the all–sky imagers.

The self–supervised model was first applied to obtain classifications of over 706 million THEMIS ASI images from 23 stations obtained between 2008 and 2022. The model returns the probabilities of the image belonging to six different classes, with the highest probability corresponding to the event label. To achieve meaningful event labels, that is, to avoid overrepresenting the same event similar to the analysis performed by Gillies et al. (2014), we downsampled the 3–s labels to 3 min. We then combined this data set with solar wind plasma parameters, IMF values, and geomagnetic indices obtained from the NASA OMNI database. After the data gaps were cleaned we removed data points where the probability of non–aurora labels, namely "cloudy," "moon," and "clear" were highest. We apply a minimum confidence threshold of 30% probability to the remainder of the data set to eliminate images where the event labels were selected between competing probabilities. We then merged the data set with ground magnetic field measurements from co–located magnetometers To investigate the response in the magnetometer data set to different aurora labels, we first calculated a station median value and removed these values from the measurements. We have then calculated dH, dD, dZ as running differences and removed the data points where magnetic field data was missing or was over 2000 nT after the station baselines were removed. Lastly, we calculated the time rate of change, dH/dt, dD/dt, and dZ/dt based on these data points. The resulting data set consisted of data points which had the following information:

- Probabilities from the model: arc, diffuse, discrete, event label determined from highest probability
- Station specific values: magnetic longitude, magnetic latitude, magnetic local time, timestamp of the image after downsampling
- Solar wind properties from OMNI Data set: IMF vector and magnitude, clock angle, and magnitude, solar wind velocity vector and magnitude, density, temperature, dynamic pressure, electric field, plasma beta, Mach number, and Alfvénic Mach number
- Geomagnetic indices from OMNI Data set: AE, AL, AU, Polar Cap index, SymD, SymH, AsyD, AsyH, Cross Polar Cap Potential
- Ground magnetic field values: H, D, Z, dH, dD, dZ, dH/dt, dD/dt, dZ/dt

After the aforementioned data preprocessing steps, the number of data points was effectively reduced to 263,303, yet this data set has significantly more data points than any aurora classification study in the literature to the best of our knowledge.

3. Results

Assessing the performance of a self–supervised model on a real–world task such as image classification is a challenge in its own right. The ability of a new self–supervised model to learn high–quality latent representations of an underlying data set is usually assessed by training on a data set for which ground–truth labels are available and then fitting a linear classifier to the latent representations (Chen et al., 2020). In our previous work using the SimCLR algorithm to classify the OATH data set (Johnson et al., 2021), we demonstrated with this approach that even with only a limited number of images a linear classifier trained on the latent representations learned by the SimCLR model achieved state of the art performance across several metrics. Without ground–truth labels for the much larger and more challenging THEMIS ASI database, performance cannot be assessed in the same way. Instead, we sample the prediction database and manually assess the accuracy of the model's predictions by visual

10.1029/2024JH000292

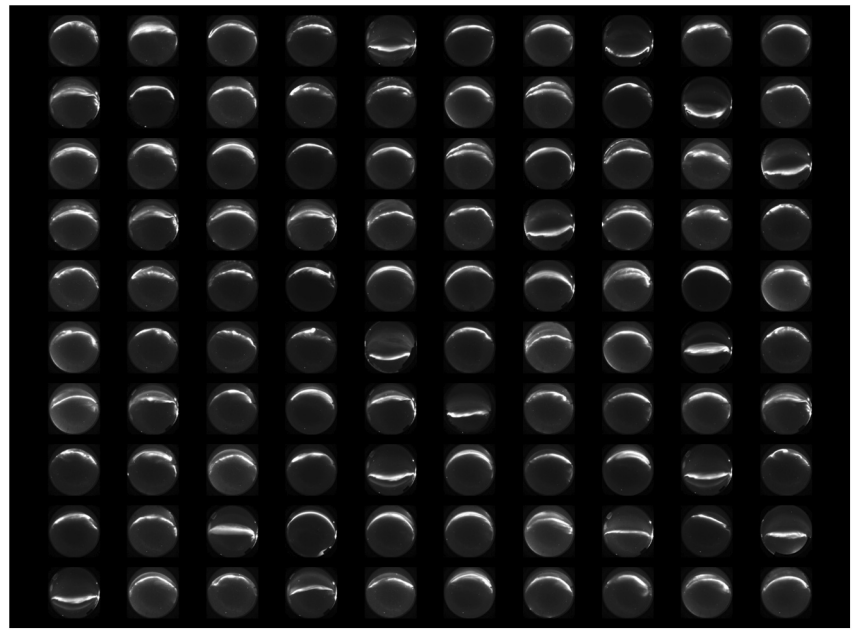


Figure 3. A random sample of images that our model classifies as "Arc" with probability greater than or equal to 40%.

inspection. To minimize the incidence of false positives, we consider only instances where the model assigns a probability greater than some predetermined threshold when assessing a classification; empirically, setting a threshold of 40% results in an average precision across classes approaching 95%. Note that it is not unusual for the model to assign equal or near–equal probabilities across classes. This may occur when a unique classification is not possible to provide; for example, if an image contains multiple phenomena such as an arc along with bright moonlight or diffuse aurora. This also happens with images that are corrupted or exhibit some out of distribution phenomenon such as a prominent human–made object in the frame or a dominant light on the horizon that is not fully cropped during preprocessing, cf. Figures 3 and 4.

There are some types of images that the model is more prone to misclassify than others. Diffuse aurora, for example, may be confused with cloudy conditions and vice versa. Arc is occasionally classified as discrete aurora,

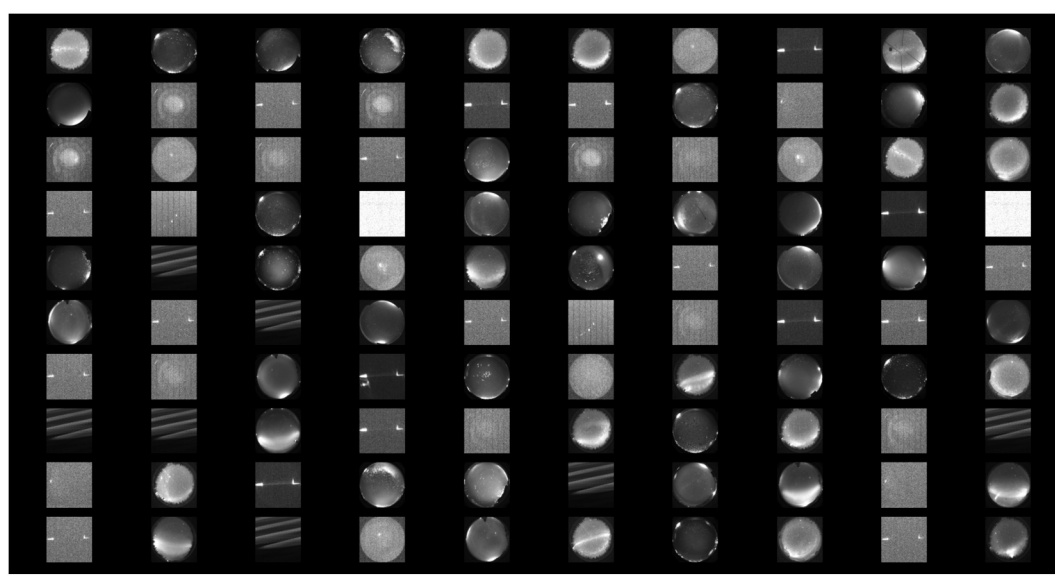


Figure 4. A collection of images that our model assigns equal probabilities of 16.6667% to each class.

29935210, 2024, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024JH000292 by Nasa Goddard Space Flight, Wiley Online Library on [10/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

29935210, 2024, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024JH000292 by Nasa Goddard Space Flight, Wiley Online Library on [10/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1
Data Preprocessing Steps for the Statistical Analysis

Step	Number of remaining data points
All labels	706,035,439
Downsampling to 3 min	11,297,325
Data merged with OMNI data points	8,266,981
Cloud, clear, and moon labels removed	1,254,693
Confidence threshold applied	1,107,828
Final data set merged with GMAG data points	861,266
Erroneous H, D, Z values removed	257,333

Note. Data preprocessing steps and the number of data points remaining after each operation.

but discrete aurora was never misclassified as arc. This behavior was less prominent when the median label for an interval was selected, therefore we have included the downsampling of the data set as a preprocessing step for statistical analysis.

The labeled data set consisted of over 706 million images as shown in the first row of Table 1. The distribution of labels post-downsampling and post-preprocessing is shown in Table 2. A summary plot of the predictions for the images from 23 stations (no data from Ekati, NT) can be seen in Figure 5 before the last two steps are applied. Each line displays the model classification for the image timestamps with colors representing the labels. The wide range of white spaces shows the summer months where classification is not possible due to a lack of ASI data. The visual representation of 1,107,828 labels for this large data set could be misleading, therefore readers are advised to take potential overlap of data points into account when interpreting the results.

With the ability to classify such a large data set, the dependence of different aurora types to external drivers, such as the interplanetary magnetic field and solar wind plasma parameters, can be further analyzed. These parameters were selected based on well-known relationships demonstrated in the literature ((Borovsky, 2021) and references therein). Figures 6a–6o shows the distribution of aurora labels with different parameters. The distributions of labels with respect to IMF components B_X and B_Y do not show a strong dependence, however, there is a clear preference for negative IMF B_Z for all three types of aurora. Figure 6f shows the peak occurrence rates of diffuse, discrete, and arc aurora as 00–02 MLT, 22–00 MLT, and 20–22 MLT respectively. For most other derived solar wind quantities and geomagnetic indices, the occurrence rates fall as there are less representative values for more intense conditions. B_X, B_Y, B_Z, MLT, E_{SW}, and CPCP exhibit normal distributions while the behaviors are skewed for B_{total}, V_X, proton density, temperature, flow pressure, Alfvénic Mach number, symH and AE values. The skewness for these distributions is most likely due to more extreme events being rare, however, it is noteworthy that the discrete labels are more comparable to diffuse labels during higher geomagnetic activity. The mean and standard deviations for these parameters are given in detail in Table 3.

Table 2
Distribution of Labels

Label	Post–Downsampling	Post–Preprocessing
All labels	11,297,325	257,333
Arc	68,239	376
Diffuse	1,413,714	187,484
Discrete	186,203	69,473
Cloudy	5,999,127	0
Moon	2,285,163	0
Clear	1,344,879	0

Note. Number of observations by label after downsampling and after removing non–auroral labels, merging with GMAG data, thresholding at 30%, and removing erroneous H, D, Z values.

To understand how the measured ground magnetic field varies with respect to auroral activity, we have used the co-located ground magnetometer measurements (H, D, Z) for corresponding stations. Figures 7a–7c show the distribution of different labels with respect to H, D, and Z components of magnetic field perturbations. Labels diffuse and discrete exhibit a normal distribution, while arc labels correspond to smaller values. Figures 7d–7f show the time rate of change of geomagnetic perturbations. It is noteworthy that even though there are more diffuse labels associated with higher geomagnetic perturbations, there are more discrete labels corresponding to higher values of magnetic field rate of change.

Finally, we investigated how the time rate of change of these perturbations differed with respect to geomagnetic activity and MLT. Figure 8 shows the MLT distribution of the rate of change with respect to symH top, a–c and AE bottom, d–f indices. It is important to note that symH is better suited for interpreting the global geoeffectiveness, while AE is widely used for

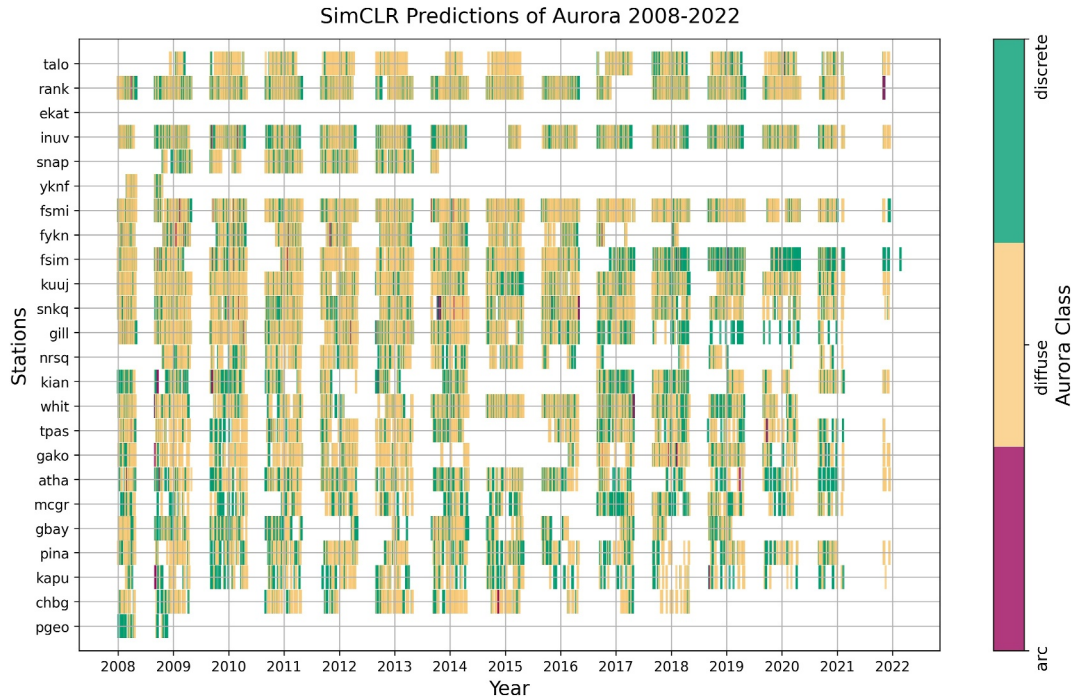


Figure 5. Distribution of predicted class for each THEMIS ASI station with respect to time between 2008 and 2022. Green color denotes discrete, yellow color denotes diffuse, and pink color denotes arc aurora labels. The stations are ordered based on magnetic latitude with the TALO station having the highest and PGEO having the lowest magnetic latitude.

interpreting geomagnetic activity at high latitudes. Figures 8a, 8b, 8d, and 8e show diffuse labels being more dominant in the post–midnight, while discrete labels are more prominent in the pre–midnight sector during intervals of increased symH and AE. As the AE drops arc aurora labels are more likely to be seen around 00 MLT.

4. Discussion

These results are complementary to other aurora classification studies as we attempt to label and study a data set spanning 12 years from 23 different stations. Even with the strict data quality flags we have applied, the resulting data set consists of 263,303 data points, which is one order of magnitude larger than any other aurora data set in the literature.

The labels used in this analysis are those assigned to the OATH data set (L. B. Clausen & Nickisch, 2018a). They represent broad auroral categories but are quite coarse, and some interesting phenomena such as auroral beads or omega bands are not identifiable with these labels. One of the primary underlying motivations for this work is to let the model learn latent representations that in the future can be used to algorithmically obtain more meaningful, fine–grained labels without necessitating manual annotation. This is a subject of our ongoing and future work.

As a result of our study, we have found that diffuse aurora occurrence peaks at 00–02 MLT, discrete aurora occurrence peaks at 22–00 MLT, and arc occurrence peaks at 20–22 MLT, respectively as shown in 6 g. These results are in agreement with Nanjo et al. (2022) who found that discrete and arc aurora peaked at approximately 23 MLT and diffuse aurora occurrence peak was at approximately 1 MLT using Tromsø all–sky camera images.

The behavior diffuse aurora exhibits with respect to MLT and high geomagnetic activity is noteworthy as the substorm expansion onset and expansion phases generally lead to discrete auroral structures like poleward boundary intensifications, streamers, unstable arcs, arc expansions and westward traveling surge in the pre–midnight sector (Akasofu, 1964). Auroral structures associated with diffuse aurora, like omega bands, generally occur as dawnward propagating structures occurring in the post–midnight sector. One immediate sanity check is to re–evaluate the model labels, especially to investigate whether the model assigns similar probabilities for discrete and diffuse sequentially. After investigating how many discrete labels exist within 5 min of a diffuse label in the data set, we have found that around 90% of diffuse events had no association with discrete labels for

29935210, 2024, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024JH000292 by Nasa Goddard Space Flight, Wiley Online Library on [10/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

10.1029/2024JH000292



Figure 6. The histograms displaying the predicted aurora classes with respect to (a) IMF B_X, (b) IMF B_Y, (c) IMF B_Z, (d) B_{Total}, (e) IMF Clock angle, (f) MLT, (g) solar wind velocity, (h) proton density, (i) temperature, (j) dynamic pressure, (k) electric field, (l) Alfvénic Mach number, (m) Sym H index, (n) calculated CPCP, and (o) AE index. Note that the light–green color here and in Figure 7 that is not explained in the legend occurs where discrete and diffuse counts overlap, and does not indicate a separate class.

29935210, 2024, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024JH000292 by Nasa Goddard Space Flight, Wiley Online Library on [10/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

29935210, 2024, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024JH000292 by Nasa Goddard Space Flight, Wiley Online Library on [10/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 3
Statistics

Parameter	Mean			Std		
Labels	Arc	Diffuse	Discrete	Arc	Diffuse	Discrete
Bx [nT]	0.26	0.04	0.23	3.12	3.71	3.65
By [nT]	−0.71	−0.12	−0.37	3.43	4.52	4.35
Bz [nT]	−0.58	−0.73	−0.98	3.13	3.76	3.65
Btotal [nT]	5.3	6.31	6.03	2.23	3.31	3.2
Clock Angle [deg]	−2.27	−1.06	−0.94	58	58	58
Vx [km/s]	−416	−453	−446	99.16	107	107
Np (cm^{-3})	5.7	6.6	6.35	4.74	5.81	5.07
Temperature [K]	85,276	112,820	106,845	76,192	99,840	100,688
Pressure [nPa]	1.81	2.48	2.31	1.34	2.05	1.67
Electric Field [mV/m]	0.28	0.43	0.55	1.31	1.76	1.7
Alfvénic Mach Number	9.91	9.71	9.71	5.87	4.84	4.7
SymH [nT]	−13	−19.32	−19.89	17.12	20.38	21.47
CPCP [kV]	18	21.62	20.97	14.13	17.08	16.8
AE [nT]	160.38	273.68	263.6	160.01	251.1	240.87

Note. The mean and standard deviation for the distributions of arc, diffuse, and discrete labels with respect to IMF B_X, IMF B_Y, IMF B_Z, B_{Total}, IMF Clock angle, solar wind velocity in x direction, proton density, temperature, dynamic pressure, electric field, Alfvénic Mach number, Sym H index, calculated CPCP, AE index.

selected stations prior to downsampling, this number is over 50% for the downsampled and combined data set that was used in the analysis. Therefore, the large time rate of change in magnetic perturbations in the pre–midnight sector during diffuse aurora could be attributed to structures like pulsating aurora which is associated with precipitating radiation belt electrons (Rankin et al., 2021). Although this behavior is mostly exhibited in the post–



Figure 7. Histograms displaying the predicted aurora classes with respect to (a) H, (b) D, (c) Z, (d) $\frac{dH}{dt}$, (e) $\frac{dD}{dt}$, and (f) $\frac{dZ}{dt}$ obtained from co-located THEMIS ground magnetometers.



Figure 8. The scatter plots displaying the predicted aurora classes of diffuse, discrete, and arc aurora with respect to MLT and SymH (top, a–c), and AE (bottom, d–f). The colors and point sizes are determined based on the amplitude of dH/dt. The point sizes are same for diffuse and discrete labels, and different for arc labels.

midnight sector, factors such as the IMF BY direction, dipole tilt, and auroral oval size could play a role in the high time rate of change during diffuse aurora in post–midnight sector (Elhawary et al., 2022; Partamies et al., 2017; Tesema et al., 2020).

5. Conclusions

In this work we used a self–supervised learning algorithm to obtain latent representations for the THEMIS all–sky imager database that we then use to classify the auroral phenomena observed in the database. We utilize data from 23 stations between 2008 and 2022, resulting in the labeling of over 706 million images. Our model appears to obtain highly accurate classifications, though we caution that assessing the performance of self–supervised models such as this one on tasks like image classification is challenging. We have analyzed the occurrence rate characteristics for these labels with respect to magnetic local time and found that they agree well with other classification studies in the literature, confirming our understanding of the substorm aurora. In future work we intend to develop additional methods to rigorously validate model performance.

We further investigated the relationship between auroral labels and solar wind parameters, interplanetary magnetic field, and geomagnetic indices. The results showed mostly normal and skewed distributions which were influenced by the occurrence rates of solar wind observations. The investigations of auroral labels with respect to ground based magnetometer signatures showed that both diffuse and discrete labels were associated with a high rate of change in ground magnetic fields. A particularly interesting finding was that during elevated levels of geomagnetic activity, diffuse labels were associated with pre–midnight sector observations of high rates of change in the ground magnetic field. This result shows that the labeled database could reveal further insight into auroral dynamics and provide a large enough sample set for future statistical studies. The labeled data set is open to the community and can be accessed via (Johnson et al., 2024b).

29935210, 2024, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024JH000292 by Nasa Goddard Space Flight, Wiley Online Library on [10/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Data Availability Statement

THEMIS ASI and GMAG data can be found at (THEMIS, 2024). OATH data can be found at (L. B. Clausen & Nickisch, 2018b). OMNI data can be found at (Papitashvili & King, 2020). The software written and used for this work is available at (Johnson et al., 2024a). The classification data produced by the models is available at (Johnson et al., 2024b).

Acknowledgments
This work was supported by National Aeronautics and Space Administration award 80NSSC22K0642 and National Science Foundation award OIA—1920965. We acknowledge NASA contract NAS5–02099 and V. Angelopoulos for use of data from the THEMIS Mission. Specifically: S. Mende and E. Donovan for use of the ASI data, the CSA for logistical support in fielding and data retrieval from the GBO stations, and NSF for support of GIMNAST through grant AGS-1004736. We thank Clausen et al. for preparing the OATH data set. Any views expressed in this work are not endorsed by the sponsors.

References

Akasofu, S.-I. (1964). The development of the auroral substorm. *Planetary and Space Science*, *12*(4), 273–282. https://doi.org/10.1016/0032-0633(64)90151-5

Akasofu, S.-I. (1981). Energy coupling between the solar wind and the magnetosphere. *Space Science Reviews*, *28*(2), 121–190. https://doi.org/10.1007/bf00218810

Borovsky, J. E. (2021). Is our understanding of solar-wind/magnetosphere coupling satisfactory? *Frontiers in Astronomy and Space Sciences*, *8*, 634073. https://doi.org/10.3389/fspas.2021.634073

Caron, M., Misra, I., Mairal, J., Goyal, P., Bojanowski, P., & Joulin, A. (2020). Unsupervised learning of visual features by contrasting cluster assignments. *arXiv preprint arXiv:2006.09882*.

Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020). A simple framework for contrastive learning of visual representations. In *International Conference on Machine Learning* (pp. 1597–1607).

Clausen, L., Milan, S. E., & Grocott, A. (2014). Thermospheric density perturbations in response to substorms. *Journal of Geophysical Research: Space Physics*, *119*(6), 4441–4455. https://doi.org/10.1002/2014ja019837

Clausen, L. B., & Nickisch, H. (2018a). Automatic classification of auroral images from the Oslo auroral themis (oath) data set using machine learning. *Journal of Geophysical Research: Space Physics*, *123*(7), 5640–5647. https://doi.org/10.1029/2018ja025274

Clausen, L. B., & Nickisch, H. (2018b). Oslo aurora THEMIS (OATH) training dataset. Retrieved from http://tid.uio.no/plasma/oath/

Dai, Z., Cai, B., Lin, Y., & Chen, J. (2021). Up-detr: Unsupervised pre-training for object detection with transformers. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 1601–1610).

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. *Cvpr09*.

Donovan, E., Mende, S., Jackel, B., Frey, H., Syrjäsuo, M., Voronkov, I., et al. (2006). The Themis all-sky imaging array—System design and initial results from the prototype imager. *Journal of Atmospheric and Solar-Terrestrial Physics*, *68*(13), 1472–1487. https://doi.org/10.1016/j.jastp.2005.03.027

Elhawary, R., Laundal, K. M., Reistad, J. P., & Hatch, S. M. (2022). Possible ionospheric influence on substorm onset location. *Geophysical Research Letters*, *49*(4), e96691. https://doi.org/10.1029/2021GL096691

Geirhos, R., Rubisch, P., Michaelis, C., Bethge, M., Wichmann, F. A., & Brendel, W. (2018). Imagenet-trained CNNS are biased towards texture; increasing shape bias improves accuracy and robustness. *arXiv preprint arXiv:1811.12231*.

Gillies, D., Knudsen, D., Donovan, E., Spanswick, E., Hansen, C., Keating, D., & Erion, S. (2014). A survey of quiet auroral arc orientation and the effects of the interplanetary magnetic field. *Journal of Geophysical Research: Space Physics*, *119*(4), 2550–2562. https://doi.org/10.1002/2013ja019469

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770–778).

Johnson, J. W., Hari, S., Hampton, D., Connor, H. K., & Keesee, A. (2021). A contrastive learning approach to auroral identification and classification. In *2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA)* (pp. 772–777).

Johnson, J. W., Öztürk, D. S., Hampton, D., Connor, H. K., Blandin, M., & Keesee, A. (2024a). Automatic detection and classification of aurora in Themis all-sky images. https://doi.org/10.5281/zenodo.11397579

Johnson, J. W., Öztürk, D. S., Hampton, D., Connor, H. K., Blandin, M., & Keesee, A. (2024b). THEMIS all-sky image classifications [Dataset]. https://doi.org/10.5061/dryad.sbcc2frft

Kim, J., Hur, Y., Park, S., Yang, E., Hwang, S. J., & Shin, J. (2020). Distribution aligning refinery of pseudo-label for imbalanced semi-supervised learning. *Advances in Neural Information Processing Systems*, *33*, 14567–14579.

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, *25*.

Kvammen, A., Wickstrøm, K., McKay, D., & Partamies, N. (2020). Auroral image classification with deep neural networks. *Journal of Geophysical Research: Space Physics*, *125*(10), e2020JA027808. https://doi.org/10.1029/2020ja027808

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278–2324. https://doi.org/10.1109/5.726791

Nanjo, S., Nozawa, S., Yamamoto, M., Kawabata, T., Johnsen, M. G., Tsuda, T. T., & Hosokawa, K. (2022). An automated auroral detection system using deep learning: Real-time operation in Tromsø, Norway. *Scientific Reports*, *12*(1), 8038. https://doi.org/10.1038/s41598-022-11686-8

NASA science mission directorate (SMD) 2014 science plan. (n.d.). Retrieved from https://science.nasa.gov/about-us/science-strategy

National Research Council. (2013). *Solar and space physics: A science for a technological society*. The National Academies Press. https://doi.org/10.17226/13060

Newell, P., Sotirelis, T., & Wing, S. (2009). Diffuse, monoenergetic, and broadband aurora: The global precipitation budget. *Journal of Geophysical Research*, *114*(A9). https://doi.org/10.1029/2009JA014326

Oord, A. V. D., Li, Y., & Vinyals, O. (2018). Representation learning with contrastive predictive coding. *arXiv preprint arXiv:1807.03748*.

Papitashvili, N. E., & King, J. H. (2020). OMNI 1-min data [Dataset]. *NASA Space Physics Data Facility*. https://doi.org/10.48322/45bb-8792

Partamies, N., Dol, B., Teissier, V., Juusola, L., Syrjäsuo, M., & Mulders, H. (2024). Auroral breakup detection in all-sky images by unsupervised learning. *Annales Geophysicae*, *42*(1), 103–115. https://doi.org/10.5194/angeo-42-103-2024

Partamies, N., Whiter, D., Kadokura, A., Kauristie, K., Nesse Tyssøy, H., Massetti, S., et al. (2017). Occurrence and average behavior of pulsating aurora. *Journal of Geophysical Research (Space Physics)*, *122*(5), 5606–5618. https://doi.org/10.1002/2017JA024039

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., et al. (2019). Pytorch: An imperative style, high-performance deep learning library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d' Alché-Buc, E. Fox, & R. Garnett (Eds.), *Advances in Neural Information*

29935210, 2024, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024JH000292 by Nasa Goddard Space Flight, Wiley Online Library on [10/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Processing Systems 32 (pp. 8024–8035). Curran Associates, Inc.. Retrieved from http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf

Rankin, R., Gillies, D. M., & Degeling, A. W. (2021). On the relationship between shear alfvén waves, auroral electron acceleration, and field line resonances. *Space Science Reviews*, *217*(4), 60. https://doi.org/10.1007/s11214-021-00830-x

Rao, J., Partamies, N., Amariutei, O., Syrjäsuo, M., & van de Sande, K. E. (2014). Automatic auroral detection in color all-sky camera images. *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *7*(12), 4717–4725. https://doi.org/10.1109/jstars.2014.2321433

Sado, P., Clausen, L. B. N., Miloch, W. J., & Nickisch, H. (2022). Transfer learning aurora image classification and magnetic disturbance evaluation. *Journal of Geophysical Research: Space Physics*, *127*(1), e2021JA029683. https://doi.org/10.1029/2021ja029683

Sado, P., Clausen, L. B. N., Miloch, W. J., & Nickisch, H. (2023). Substorm onset prediction using machine learning classified auroral images. *Space Weather*, *21*(2), e2022SW003300. https://doi.org/10.1029/2022sw003300

Sharif Razavian, A., Azizpour, H., Sullivan, J., & Carlsson, S. (2014a). Cnn features off-the-shelf: An astounding baseline for recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (pp. 806–813).

Sharif Razavian, A., Azizpour, H., Sullivan, J., & Carlsson, S. (2014b). Cnn features off-the-shelf: An astounding baseline for recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*.

Sohn, K. (2016). Improved deep metric learning with multi-class n-pair loss objective. In *Proceedings of the 30th International Conference on Neural Information Processing Systems* (pp. 1857–1865).

Syrjäsuo, M., & Donovan, E. (2002). Analysis of auroral images: Detection and tracking. *Geophysica*, *38*(1–2), 3–14.

Syrjäsuo, M., & Donovan, E. (2004). Diurnal auroral occurrence statistics obtained via machine vision. *Annales Geophysicae*, *22*(4), 1103–1113. https://doi.org/10.5194/angeo-22-1103-2004

Syrjäsuo, M., Donovan, E., Qin, X., & Yang, Y. (2007). Automatic classification of auroral images in substorm studies. In *8th International Conference on Substorms (ICS8)* (pp. 309–313).

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., et al. (2015). Going deeper with convolutions. In *Proceedings of the Ieee Conference on Computer Vision and Pattern Recognition* (pp. 1–9).

Tesema, F., Partamies, N., Nesse Tyssøy, H., & McKay, D. (2020). Observations of precipitation energies during different types of pulsating aurora. *Annales Geophysicae*, *38*(6), 1191–1202. https://doi.org/10.5194/angeo-38-1191-2020

THEMIS. (2024). Time history of events and macroscale interactions during substorms. Retrieved from https://themis.ssl.berkeley.edu/index.shtml

Wu, Z., Xiong, Y., Yu, S. X., & Lin, D. (2018). Unsupervised feature learning via non-parametric instance discrimination. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3733–3742).

Yang, Q., Liang, J., Hu, Z., & Zhao, H. (2012). Auroral sequence representation and classification using hidden Markov models. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(12), 5049–5060. https://doi.org/10.1109/tgrs.2012.2195667

Yang, X., Song, Z., King, I., & Xu, Z. (2022). A survey on deep semi-supervised learning. *IEEE Transactions on Knowledge and Data Engineering*, *35*(9), 8934–8954. https://doi.org/10.1109/tkde.2022.3220219

29935210, 2024, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024JH000292 by Nasa Goddard Space Flight, Wiley Online Library on [10/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License